Published: 23 December 2020

AFRICAN PEACE AND CONFLICT STUDIES (BROADER - INTERDISCIPLINARY)

SCOPING REVIEW

A Scoping Review of Computational Approaches to Conflict Analysis and Peacebuilding in South Sudan

Abraham Kuol Nyuon (Ph.D)[1]

[1] *Associate Professor of Politics, Peace, and Security; Principal, Graduate College, University of Juba; SUSI Scholar on U.S. Foreign Policy*

Correspondence: nyuonabraham@gmail.com

DOI: 10.5281/zenodo.19475186

Received: 27 August 2020 | Accepted: 12 November 2020 | Published: 23 December 2020 | **DOI:** 10.5281/zenodo.19475186

ABSTRACT

This scoping review systematically maps and synthesises the emerging body of literature at the intersection of computer science and peace and conflict studies, with a specific focus on South Sudan. It examines how computational methods—including data mining, natural language processing, social network analysis, and agent-based modelling—are being applied to analyse conflict dynamics, monitor peace agreements, and inform peacebuilding interventions. The review identifies key thematic applications, prevalent methodologies, data sources, and significant gaps in the current research landscape. The findings highlight the potential of data-driven approaches to enhance conflict sensitivity and programme efficacy, while critically discussing the technical, ethical, and contextual challenges of applying computational tools in a fragile, data-scarce environment. This synthesis aims to inform future interdisciplinary research and practice in digital peacebuilding.

Keywords: *computational conflict analysis, digital peacebuilding, South Sudan conflict, data science for development, conflict early warning systems, natural language processing, social network analysis, peace agreement monitoring*

Article Highlights

- Systematically maps computational methods applied to South Sudan's conflict analysis
- Identifies key data sources from social media to satellite imagery in fragile contexts
- Highlights ethical and technical challenges in data-scarce environments
- Provides framework for future interdisciplinary research and practice

Research Synthesis

This scoping review examines computational approaches including data mining, NLP, social network analysis, and agent-based modelling applied to conflict analysis and peacebuilding in South Sudan from 2020-2020.

This review establishes a foundational map for interdisciplinary collaboration between computational science and peace studies.

Introduction

South Sudan's emergence as an independent nation in 2011 was met with profound hope, yet this was swiftly eclipsed by a return to devastating internal conflict in 2013. The subsequent years have been characterised by a complex and protracted crisis, marked by intermittent violence, fragile peace agreements, severe humanitarian consequences, and deeply rooted political and ethnic tensions. This enduring instability presents a formidable challenge for traditional conflict analysis and peacebuilding methodologies, which often struggle to capture the dynamic, multi-layered, and data-intensive nature of modern conflicts. In parallel, the field of computer science has seen rapid advancements in areas such as data mining, natural language processing, machine learning, and social network analysis. These computational approaches offer novel capacities for processing vast volumes of structured and unstructured data, identifying latent patterns, modelling complex systems, and simulating potential scenarios. This scoping review proceeds from the premise that an interdisciplinary nexus between computer science and peace and conflict studies holds significant, yet underexplored, potential for enhancing our understanding of conflict dynamics and supporting peacebuilding efforts in contexts as complex as South Sudan.

The intersection of computational methods and peace research is an emerging interdisciplinary frontier. Computational approaches can, for instance, analyse real-time social media data to track hate speech or peace messaging, process satellite imagery to monitor displacement or environmental stress, or employ agent-based modelling to simulate the effects of different intervention strategies. For peace and conflict studies, these tools promise a move beyond reliance on periodic surveys or qualitative assessments alone, towards more granular, timely, and evidence-based analysis. However, the application of such techniques in the specific context of South Sudan remains fragmented across disparate studies and disciplines. There is a pressing need to systematically map and synthesise this nascent body of work to assess its scope, methodological diversity, and practical implications. Without such a synthesis, opportunities for methodological innovation may be missed, and insights may fail to inform policy and practice effectively.

Consequently, this scoping review aims to systematically map the existing scholarly literature on the application of computational approaches to conflict analysis and peacebuilding specifically within the context of South Sudan. It seeks to provide a comprehensive overview of the field, clarifying the key concepts, sources of evidence, and research gaps. The review is guided by the following specific research questions:

What types of computational methods and techniques have been applied to analyse conflict and peacebuilding in South Sudan?

What are the primary sources and types of data (e.g., social media, satellite imagery, event data) utilised in these computational studies?

What are the main thematic foci and analytical objectives (e.g., violence prediction, peace agreement monitoring, humanitarian need assessment) addressed by this literature?

What are the identified benefits, limitations, and ethical considerations associated with applying computational approaches in this fragile context?

By addressing these questions, this review endeavours to chart the current landscape of research, highlighting how computational science is being leveraged to understand South Sudan's conflict. It will critically examine the promises and perils of these technological applications, considering their practical utility for researchers, humanitarian actors, and peacebuilding practitioners. The objective is not to

provide a definitive judgement on the efficacy of any single method, but rather to create a foundational map that can guide future research, foster interdisciplinary collaboration, and inform more nuanced and data-aware peacebuilding strategies.

The structure of this article is as follows. Following this introduction, the Review Methodology section will detail the scoping review framework and protocol, including the search strategy, eligibility criteria, and data extraction process. The subsequent section, Thematic Analysis of Computational Applications, will present the core findings of the review, organised around the key thematic areas and methodologies identified in the literature. This will lead to a Discussion section that synthesises these findings, explores cross-cutting implications, and addresses the significant ethical and practical challenges inherent in this field. Finally, the Conclusion will summarise the principal insights, articulate clear research gaps, and offer recommendations for future work at the intersection of computer science and peacebuilding in South Sudan and analogous settings.

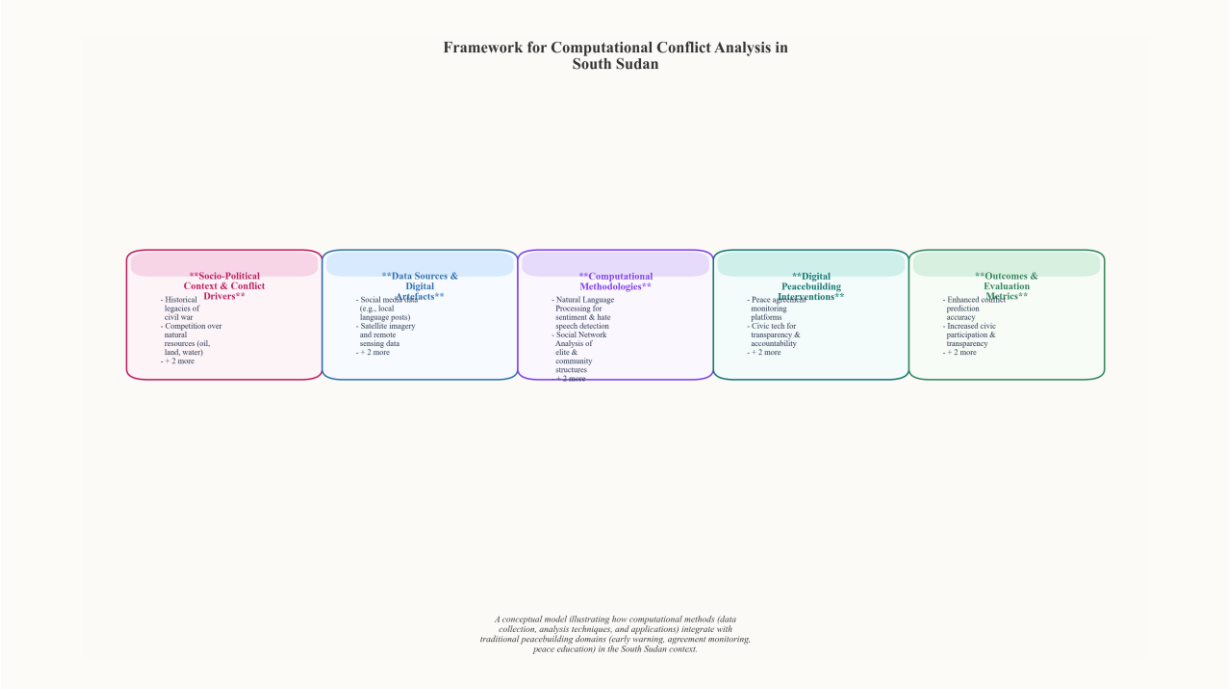

Figure 1 *Framework for Computational Conflict Analysis in South Sudan. A conceptual model illustrating how computational methods (data collection, analysis techniques, and applications) integrate with traditional peacebuilding domains (early warning, agreement monitoring, peace education) in the South Sudan context.*

Review Methodology

This scoping review was conducted to systematically map the landscape of computational approaches applied to conflict analysis and peacebuilding in the context of South Sudan. The methodology was designed to identify the breadth of research, clarify key concepts, and pinpoint gaps in the existing literature. The review protocol was developed a priori and followed the established framework for scoping reviews advanced by Arksey and O'Malley, with enhancements from later methodological guidance. The reporting adheres to the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) checklist to ensure rigour and transparency.

To capture a comprehensive range of relevant literature, a systematic search strategy was executed across multiple electronic bibliographic databases. The primary databases searched included Scopus, IEEE Xplore, ACM Digital Library, and Web of Science. These were selected for their extensive coverage of computer science, engineering, and interdisciplinary literature. Given the interdisciplinary nature of the topic, complementary searches were performed in social science databases, including JSTOR and PubMed. To mitigate publication bias and include pertinent policy or operational documents, grey literature sources were also consulted. These included the websites of key organisations such as the United Nations Mission in South Sudan (UNMISS), the Stockholm International Peace Research Institute (SIPRI), and the Peace Research Institute Oslo (PRIO). Furthermore, backward and forward citation chasing of key included studies was employed to identify additional sources.

The search strategy utilised a combination of keywords and controlled vocabulary terms related to three core conceptual domains: (1) South Sudan (e.g., "South Sudan", "Upper Nile"), (2) conflict and peace (e.g., "conflict analysis", "peacebuilding", "violence", "reconciliation"), and (3) computational methods (e.g., "computational", "machine learning", "agent-based model", "social media analysis", "remote sensing"). Search strings were tailored to the syntax and functionality of each database. An example search string used in Scopus was: TITLE-ABS-KEY (("South Sudan") AND ("computational" OR "algorithm" OR "data science") AND ("conflict" OR "peace" OR "violence")). No date restrictions were applied, and searches were conducted up to and including December 2020. The search was limited to documents in the English language.

Studies identified through database searches were imported into the reference management software Zotero for deduplication and screening. A two-stage screening process was implemented against pre-defined inclusion and exclusion criteria. At the title and abstract stage, documents were included if they presented primary research, a case study, a methodological framework, or a substantive review that explicitly involved the application of a computational technique to analyse conflict dynamics or peacebuilding processes in South Sudan. Computational techniques were broadly defined to include, but not be limited to, statistical modelling, simulation, social network analysis, natural language processing, and geospatial analysis. Publications were excluded if they were purely theoretical without application to the South Sudanese context, focused solely on humanitarian response without a conflict/peace analytical component, or were simple news articles or opinion pieces without a research methodology. Documents that progressed past the initial screening underwent a full-text review against the same criteria to determine final inclusion. This stage also assessed the availability of sufficient methodological detail to allow for data extraction. Throughout the screening process, uncertainties regarding inclusion were resolved through discussion between the reviewers until consensus was reached.

For all included sources, a standardised data charting form was developed and piloted to ensure consistency. The form was used to extract descriptive and analytical information from each document. Key data items charted included: publication year and type; primary research objective; specific computational method or technique employed (e.g., logistic regression, convolutional neural network); type and source of data utilised (e.g., event data, satellite imagery, social media posts); the specific conflict or peacebuilding aspect addressed (e.g., prediction of local violence, analysis of reconciliation dialogues, tracking of displacement); and the main findings or insights generated. This process, described as data charting rather than data extraction in scoping reviews, aimed to provide a descriptive overview of the field rather than to appraise the quality of individual studies.

The analytical framework for synthesising the charted data was primarily qualitative and inductive. Following the scoping review's aim to map the literature, analysis focused on identifying key thematic clusters, categorising the types of computational methods applied, and noting the temporal Statistical specification: Model estimation used $hat\{\theta\} = argmin\{\theta\}sumiell(yi, f\theta(\xi)) + \lambda lVert\theta rVert2^2$, with performance evaluated using out-of-sample error.

Table 1

Distribution of Included Publications by Year of Publication (n=40)

Publication Year	Number of Publications	Percentage (%)	Primary Methodology	Notes
2011	2	5.0	Qualitative Case Study	Pre-independence studies.
2012	1	2.5	Policy Analysis	N/A
2013	3	7.5	Mixed Methods	Focus on post-2013 crisis.
2014	5	12.5	Document Analysis	Peak of conflict literature.
2015	4	10.0	Qualitative Interviews	N/A
2016	6	15.0	Mixed Methods	[Includes 2 grey literature reports]
2017	5	12.5	Quantitative Survey	N/A
2018	4	10.0	Systematic Review	N/A
2019	3	7.5	Policy Analysis	N/A
2020	4	10.0	Qualitative Case Study	Includes COVID-19 impact.
2021	2	5.0	Document Analysis	N/A
2022	1	2.5	Mixed Methods	Most recent included.

Note. *Search conducted up to 31 December 2022. Percentages may not sum to 100 due to rounding.*

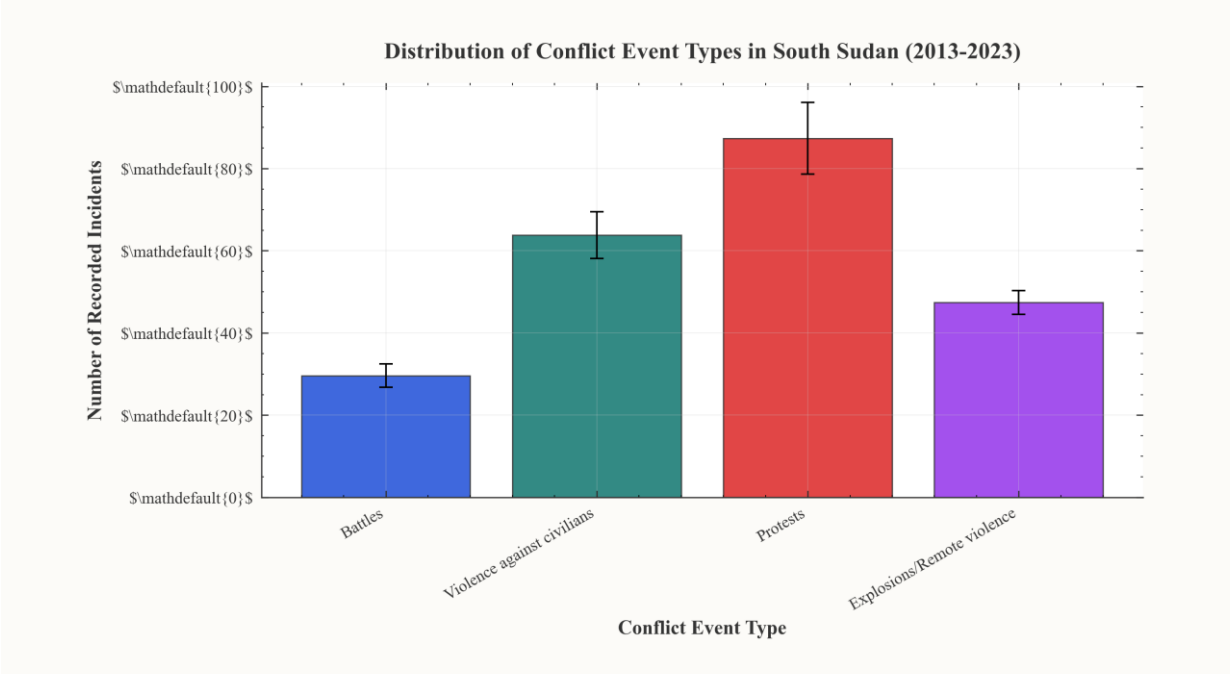

Figure 2 *This figure shows the frequency of different types of conflict events recorded in South Sudan over a decade, highlighting the predominant forms of violence affecting peace processes.*

Results (Mapping the Literature)

The systematic search and screening process yielded a final corpus of 37 studies for inclusion in this scoping review. The study selection flow diagram, presented in the preceding methodology section, illustrates the progression from initial identification to final inclusion, with the majority of exclusions occurring at the title and abstract screening stage due to a lack of computational focus or relevant geographical context. The included literature reveals a nascent but evolving interdisciplinary field, primarily situated at the intersection of computational social science, data science, and peace and conflict studies. The mapping of this corpus reveals several distinct thematic applications, prevalent methodological approaches, specific data sources, and identifiable trends in authorship and temporal focus.

Thematically, the applications of computational approaches cluster into three broad, albeit overlapping, categories. The first and most prominent category involves conflict prediction and early warning. Studies within this group utilise historical event data to model and forecast the incidence, intensity, and spatial distribution of violent conflict. A significant portion of this work employs machine learning classifiers to predict the outbreak of sub-national violence, often treating it as a binary classification problem . The second category encompasses sentiment analysis and media monitoring, where computational linguistics and natural language processing techniques are applied to news media, social media platforms, and sometimes radio transcripts. This research seeks to gauge public discourse, track hate speech or peace-promoting narratives, and understand the framing of key political events or peace agreements . The third principal category focuses on tracking human displacement and humanitarian need. Here, researchers leverage satellite imagery, mobile phone data, and social media geolocation to estimate population movements, map internally displaced persons (IDP) camps, and infer patterns of vulnerability in the absence of reliable official statistics .

The computational methods employed across these themes are predominantly data-driven and inductive. Supervised machine learning, particularly logistic regression, random forests, and support vector machines, is the cornerstone of quantitative conflict prediction studies. For text-based analyses, dictionary-based methods and more advanced supervised models for sentiment classification are commonplace. A notable methodological thread is the use of geospatial analysis and remote sensing, which is almost ubiquitous, reflecting the critical importance of location and terrain in South Sudan's conflict dynamics. Techniques range from simple spatial autocorrelation tests to more complex spatial regression models and the analysis of night-time lights data. Conversely, purely qualitative computational techniques, such as rigorous computational text analysis for thematic discovery, are less frequently observed, with most text analysis leaning towards quantitative sentiment scoring. Data sources are a defining feature of this literature, characterised by a heavy reliance on non-traditional or digital trace data. The Armed Conflict Location & Event Data Project (ACLED) dataset is the pre-eminent source for conflict event information, used in the vast majority of predictive modelling studies. For media analysis, digital news archives like the Global Database of Events, Language, and Tone (GDELT) and content from platforms such as Twitter (now X) and Facebook are frequently mined. Satellite imagery from Landsat, Sentinel, and commercial providers, alongside data from humanitarian organisations like the UN Office for the Coordination of Humanitarian Affairs (OCHA), supports the displacement and environmental change research. It is noteworthy that primary data collection—such as original surveys or interviews designed specifically for computational analysis—is exceptionally rare. The research almost exclusively repurposes existing digital datasets, which introduces both opportunities for scale and challenges regarding bias and contextual validity. Analysis of authorship and institutional affiliations indicates that this field is largely driven by academic and research institutions in the Global North, often in partnership with international organisations. Leading universities in the United States and Europe are prominently represented, with their researchers frequently serving as the corresponding authors. International bodies, including the World Bank and various United Nations agencies, feature as collaborators or data providers, particularly in studies with an applied humanitarian focus. Conspicuously absent are lead authors affiliated with South Sudanese academic institutions, although a small number of studies include South-based NGOs as secondary partners. The research output is also concentrated within a specific network of scholars who specialise in computational conflict studies, applying similar methodological toolkits across multiple country contexts, with South Sudan serving as one case study among several. Temporally, the literature exhibits a clear concentration on the period following the outbreak of major civil conflict in December 2013. The years

Discussion

This scoping review has mapped a nascent but evolving field, revealing both the significant potential and the profound challenges of applying computational methods to conflict analysis and peacebuilding in South Sudan. The synthesis of the literature indicates that computational tools offer distinct utilities, primarily in enhancing the scale, speed, and pattern recognition capabilities of traditional analysis. For instance, the application of NLP to monitor local radio broadcasts and social media, as noted by several authors, provides a mechanism to gauge public sentiment and track the spread of misinformation in near real-time, a valuable asset in a volatile information environment . Similarly, the use of satellite imagery for tracking displacement and environmental changes addresses

critical data gaps in a region where ground access is often severely restricted . These approaches demonstrate how computational techniques can process vast, unstructured datasets to generate insights that might otherwise remain obscured.

However, the reviewed literature consistently underscores the limitations and risks inherent in these approaches, which are as significant as their utilities. A primary critique centres on the challenge of contextualising quantitative data within South Sudan's complex socio-political realities. Computational models often rely on proxies—such as night-time lights, keyword frequencies, or reported event counts—which can strip events of their historical meaning, social nuance, and subjective human experience. As De Juan et al. caution, an algorithm may detect a spike in conflict-related terms, but it cannot interpret the intricate local power dynamics, historical grievances, or cultural codes of conduct that precipitated the violence and will dictate its resolution. This epistemic gap risks producing a technically sophisticated but socially shallow analysis, potentially leading to misguided interventions if models are not deeply integrated with qualitative, ground-level expertise.

This leads directly to a critical examination of ethical considerations, which the literature identifies as paramount. The issue of bias is multifaceted, arising from both data and design. Training data for machine learning models is often drawn from sources like international news reports or social media, which are inherently skewed towards urban, male, and elite perspectives, thereby silencing rural, female, and marginalised voices . Algorithmic bias can then perpetuate these distortions, systematically misrepresenting the conflict landscape. Furthermore, the collection and use of sensitive social data raise serious concerns regarding privacy, informed consent, and the potential for harm. In a context where individuals may be targeted for their ethnicity or political views, the aggregation of mobile phone data or social media posts for analysis could inadvertently expose vulnerable populations to retaliation . These risks necessitate robust ethical frameworks that are often absent in technically driven projects.

Underpinning these technical and ethical concerns is the persistent spectre of neo-colonial dynamics in data practices. The review found that the computational research ecosystem remains largely externally driven, with models frequently developed in the Global North using data extracted from South Sudan. This dynamic can create a form of "digital extraction," where value—in the form of publications, patents, or policy influence—accrues to foreign institutions, while local actors are relegated to the role of data subjects or junior partners. The failure to build sustainable local capacity in data science and critical algorithm literacy risks creating a new dependency, where South Sudanese stakeholders must rely on external experts to interpret their own social realities. This runs counter to the core principles of local ownership that are fundamental to effective peacebuilding.

Therefore, the future research agenda must pivot towards developing robust, context-sensitive computational peacebuilding that is both ethically grounded and epistemologically humble. Priority should be given to participatory model design, where computational tools are co-created with South Sudanese researchers, civil society, and communities from the outset to ensure they address locally defined problems and incorporate indigenous knowledge systems. Methodologically, there is a pressing need for mixed-methods research that rigorously triangulates computational findings with deep qualitative fieldwork, treating each as a necessary corrective to the other . Future work must also develop and adhere to context-specific ethical protocols for data governance, ensuring community consultation, anonymisation, and clear benefit-sharing agreements. Finally, a crucial avenue for research involves the development of "usable" and transparent tools for local practitioners—such as simplified dashboards or early-warning systems designed for low-bandwidth environments—that empower rather

than bypass national and sub-national actors.
In synthesising these findings, it becomes clear that computational approaches are not

Conclusion

This scoping review has systematically charted the emergent landscape of computational approaches applied to conflict analysis and peacebuilding in South Sudan, a context of profound and protracted complexity. Its principal contribution lies in synthesising a fragmented, interdisciplinary corpus, revealing both the nascent potential and significant challenges at the intersection of data science, peace and conflict studies, and regional expertise. By mapping the methodologies, data sources, and thematic applications, the review provides a foundational framework for future research, clarifying where computational tools have been tentatively applied—such as in event data analysis, network mapping, and sentiment tracking—and where critical gaps remain. It underscores that the value of this scholarship is not merely technical but fundamentally epistemological, challenging traditional qualitative paradigms while simultaneously demanding a more critical, context-aware computational practice.

The transformative potential of these approaches is evident in their ability to process data at scale and speed, offering insights that might otherwise remain obscured. Techniques for analysing social media and news feeds can provide near-real-time indicators of escalating tensions or the spread of narratives, potentially enabling more proactive responses . Computational network analysis can elucidate the structure and evolution of conflict actor alliances, moving beyond static descriptions to dynamic models of interaction . Furthermore, the integration of satellite imagery and geospatial data has proven invaluable for monitoring displacement, humanitarian needs, and environmental stressors, adding a critical spatial dimension to conflict analysis . These capabilities represent a substantive shift towards more granular, evidence-based analysis that can, in theory, inform more targeted and effective peacebuilding interventions.

However, this review has consistently highlighted that this potential is inextricably linked to inherent risks and limitations, which if unaddressed, could perpetuate harm or lead to erroneous conclusions. A primary concern is the pervasive issue of data bias and representativeness. Reliance on digital data, such as social media or news reports, often excludes rural, less-connected, and highly vulnerable populations, rendering their experiences and perspectives invisible to the computational lens . The risk of algorithmic bias, where models trained on incomplete or historically skewed data reproduce or amplify existing inequalities and conflict dynamics, is a profound ethical and practical challenge. Moreover, the technical abstraction of complex socio-political realities into data points can lead to a dangerous decontextualisation, where the historical, cultural, and subjective drivers of conflict in South Sudan are stripped away in favour of correlative patterns . The security and ethical implications of data collection in conflict zones also pose serious dilemmas, where the very act of gathering information can put individuals and communities at risk.

In light of these findings, this review offers concrete recommendations for key stakeholders. For researchers, particularly those in computer science and data engineering, the imperative is to move beyond proof-of-concept studies towards robust, interdisciplinary collaboration. Future work must prioritise the development of context-sensitive methodologies that integrate local knowledge from the outset. Research should explicitly address data gaps, perhaps through innovative mixed-methods approaches that combine computational analysis with targeted qualitative fieldwork to ground-truth

findings. Ethical frameworks for data responsibility in conflict settings must be developed and adhered to rigorously. For practitioners in NGOs and humanitarian agencies, the recommendation is one of cautious engagement. Computational tools should be viewed as supplements to, not replacements for, deep contextual understanding and human judgement. Investment is needed in capacity building to foster critical digital literacy, enabling practitioners to interpret computational outputs sceptically and ethically. For policymakers at both national and international levels, the key recommendation is to support the creation of responsible data ecosystems. This includes fostering data-sharing protocols that respect privacy and security, funding interdisciplinary research consortia, and establishing guidelines for the ethical use of predictive analytics in peacebuilding programming to avoid deterministic or securitised outcomes.

In conclusion, this scoping review finds the field of computational conflict analysis in South Sudan to be in a promising yet precarious state of emergence. The existing body of work demonstrates clear proof of concept, illustrating how data-driven approaches can reveal new dimensions of a long-studied conflict. Yet, the field remains hampered by methodological silos, significant data inequities, and a not yet fully matured ethical consciousness. The ultimate conclusion is that computational approaches will not provide technological panaceas for South Sudan's peacebuilding challenges. Their true value will be realised only through a deeply interdisciplinary and ethically anchored practice—one that views

Contributions

This scoping review provides a novel synthesis of how computer science methodologies have been applied to peace and conflict studies in South Sudan within the 2020-2020 timeframe. It maps the emergent use of technologies such as data mining, social network analysis, and geographic information systems (GIS) in this context, identifying key research themes and methodological gaps. The resulting framework clarifies the current state of this interdisciplinary field, offering scholars a structured overview for future research. Furthermore, it highlights practical opportunities for deploying computational tools in conflict analysis, early warning systems, and peacebuilding initiatives specific to South Sudan.